REVIEW OF
COLUMBIA RIVER TEMPERATURE ASSESSMENT: KALMAN FILTERING

INTRODUCTION

The purposes of this supplementary review are to:

(i) determine whether the Kalman filter has been used in a correct manner, i.e., in a manner consistent with published (and peer-reviewed) practice;

(ii) assess whether this filtering framework is appropriate for the given task;

(iii) consider alternatives to this framework suitable for any further such studies;

(iv) indicate what might be the desirable subjects of those future studies;

Before reporting on these matters, it will be helpful to review what has actually been achieved through applying the Kalman filter in this context.

GOAL OF EPA REPORT

Without seeking to diminish the significance of, or distract attention away from, variations in stream temperature over the entire annual cycle, let me suggest the real issue here is that of forecasting the maximum temperature over this cycle. Furthermore, if one is conservative in outlook, it may be better to over-predict than to under-predict this maximum. Re-stating this goal is important, because it has a bearing on some of the detail surrounding the way in which the filter has been used in the EPA Report.

"VALUE ADDED" THROUGH USE OF THE KALMAN FILTER

Besides the obviously highly charged political context of this problem, removing dams from the Columbia and Snake Rivers is a rather dramatic piece of policy. It would therefore seem important for all concerned to be reassured that such action is "right for the situation" and to be aware of the risks of "getting it wrong". Considerations of uncertainty and risk, then, are entirely appropriate in such a problem setting. Indeed, to have undertaken this exercise in the absence of such considerations, i.e., under the assumption of an entirely deterministic model, would itself have been an act of engaging in risk-taking (for the EPA). Use of the Kalman filter to address these issues of uncertainty is not usual, but by no means unknown.

The following is the essential role played by the filter in this study, to paraphrase (in perhaps colloquial terms):

> The world is uncertain. We know too that all models are approximations. All sources of uncertainty (approximations, omissions, errors) in the model will be subsumed under the labelof the system noise vector ($\underline{w}$). Besides estimates of the model's conventional (deterministic) parameters, we shall therefore need estimates of the variance-covariance properties of $\underline{w}$ (the matrix Σ_Q) in order to account for the manner in which the inevitable residual uncertainty attaching to the model -- even when calibrated -- is propagated forward into forecasts of future behavior (under changed conditions).

In fact, looking at the source reference of van Geer *et al.* (1991), one might go so far as to say the primary purpose of calibration in the present study is to adjust the estimates of the variance-covariance matrix of the system noise (Σ_Q), with a view to assessing its impact on the uncertainty of the forecasts.

To be clear about what is subsumed under this matrix, we have the following generic sources of uncertainty:

(i) uncertainty in the (deterministic) parameters of the model;

(ii) uncertainty in the measured input disturbances of the model, i.e., here, principally the variations in the temperature of the tributaries;

(iii) uncertainty in all other unmeasured disturbances of the factors affecting temperature (the state variable).

In addition, account must be taken of uncertainty in the system's (past) observations, as must the uncertainty in the initial state of the system, i.e., the values of the spatial distribution of temperatures at the start of the calibration period and the forecasting period (although the author does not discuss this source of uncertainty). To be complete, we should also note that there will be a "structural error", or conceptual error, in the model. The manner in which the model's state variables interact with each other and the forms of the expressions used to describe these interactions will diverge from the (unknowable) "truth". There is currently no adequate method of accounting for errors of this form. This is hardly surprising: the problem is more philosophical than technical.

Given the decisions to account for uncertainty in this problem and to account for it using the Kalman filter, lumping the uncertainty in this manner under the single quantity (Σ_Q) is a pragmatic restriction, consistent with benefitting from the relative computational economy of the linear Kalman filter, when set against the alternative of Monte Carlo simulation, say. It also avoids having to use an extended Kalman filter, which would be necessary if one were to separate out (from Σ_Q) the parameteric uncertainty of the model, i.e., item (i) of the sources of uncertainty listed above (see also Beck and Halfon, 1991). The disadvantage of working with such an aggregate measure of uncertainty is that it will foreclose on any analysis of ranking the relative importance of the various sources of uncertainty, in terms of their contributions to the uncertainty of the predictions. Knowledge of this latter would be important in subsequently setting priorities for work that would be needed in order to reduce prediction uncertainty to some acceptable level (if it were thought to be unacceptably high for the purposes of making decisions).

PROPER USE OF THE KALMAN FILTER

As far as can be determined, Yearsley has used the Kalman filter -- for the purpose of calibration -- strictly in accordance with the procedures set out in the paper by van Geer *et al.* (1991). These authors, in their turn, make reference to the covariance-matching procedure of Mehra (1972), which, in spite of its vintage, remains the most common method for calibrating the variance-covariance properties of the system noise, i.e., for assigning values to the elements of the matrix Σ_Q. For the purpose of predicting the consequences of the policy options, again the filter has been applied in a manner consistent with normal practice (Beck and Halfon, 1991). To this extent, no fault can be found with the filter's application here; technically, the analysis appears to be sound.

There are, however, a number of places where caution should be exercised in interpreting the results of the Report. These are as follows.

(i) Figures 6 through 13 show comparisons of the simulated and observed water temperatures. Although we cannot be certain, it is quite probable that the corrected, or updated, estimate of the water temperature, i.e., $\underline{T}_k(+)$ from equation (9), has been used as the "simulated" value. If this is so, it is important to bear in mind that the results of these Figures may suggest a performance of the model better than what would have been achieved in the more familiar, purely deterministic setting, wherein the model is not embedded within a Kalman filter. Close inspection of equation (9) reveals the presence of the current observation of temperature $\underline{z}_k$. The effect of updating the one-step-ahead prediction, $\underline{T}_k(-)$, is thus always to draw any erroneous such prediction back towards the observation. The updated estimates $\underline{T}_k(+)$ reflect the benefit of this correction. To the eye trained on assessing a model's performance in the deterministic setting, without this "tracking" feature, a comparison of $\underline{T}_k(+)$ with the observation $\underline{z}_k$ can be deceptive. It might therefore be desirable to ask for clarification of whether the "simulated" values of Figures 6-13 represent $\underline{T}_k(+)$ or $\underline{T}_k(-)$.

(ii) Of the three policy options assessed (business-as-usual, removal of dams, control of tributary temperatures), the removal of the dams will clearly lead to a hydraulic regime unlike that of the (post-dam) observed record. The most obvious expectation of the consequences of this is that the uncertainty attaching to the hydraulic parameters estimated through the approximations of equations (13) through (15), if not any of the other (deterministic) model parameters, will be greater for this regime than for the presently observed conditions (with the dams in place). As far as can be established, no account is taken of this greater uncertainty; the same values of Σ_Q are used in generating the confidence bounds around all three sets of predictions. Since removal of the dams -- on the basis of the current analysis -- is predicted to have a significantly beneficial impact on lowering the number and magnitude of violations of the maximum temperature constraint, more detailed consideration of this point may well be warranted. Furthermore, the potential significance of this particular source of uncertainty may make it appropriate for future analyses to be based on explicit representation of the constituent sources of uncertainty, as opposed to their being lumped under Σ_Q.

(iii) It appears that the variance-covariance matrix of the system noise (Σ_Q) has non-zero elements on its leading diagonal alone, i.e., the assumption has been made that disturbances of the stream temperature dynamics are uncorrelated (primarily in space, it would appear). The Report is largely silent on the making of this assumption, although it is a common and not unreasonable one. Nevertheless, there is no discussion of its possible consequences, which is unfortunate since these may be material to the analysis. It is fairly widely appreciated that covariance among the elementary sources of uncertainty can have a significant effect on the propagation of uncertainty. In fact, it has generally been thought that it has the effect of reducing the degree of uncertainty attaching to the forecasts (this is not always the case, however; Beck and Halfon, 1991). We may note that van Geer *et al.* (1991) provide a means of assigning values to these off-diagonal elements of Σ_Q; it does not appear to have been used in the present analysis.

(iv) Comparing Figures 14 through 21 with respectively Figures 6 through 13 of the Report, is a surprisingly confusing task. If the principal issue at stake in this study is under-prediction of the maximum (summer) temperatures, it is especially important to be comfortable with the fact that the innovations ($\underline{\nu}_k$) are consistent with the relative positions of the quantities, $\underline{T}_k(+)$ (assumed) and $\underline{z}_k$, plotted in their respective Figures. Even after considerable reflection, I have failed to reconcile -- to my satisfaction -- the two sets of Figures.

To summarize, the subject of this review is a Report on a screening analysis designed to identify issues for further study. In general, the Kalman filter has been properly used for this purpose. However, the author of the Report has not identified all of the issues worthy of more detailed scrutiny.

APPROPRIATENESS OF FILTERING FRAMEWORK

In strategic terms, as already stated, it seems appropriate for uncertainty and risk to be parts of this assessment. In tactical terms, the Kalman filter provides (literally) a first-order approximation of error propagation. On balance this would appear commensurate with a preliminary screening analysis, although it is not common to find the Kalman filter employed in a study of this kind. In general, one could say the filter is often a good technique for problem discovery and definition, but one might subsequently want to apply some other form of analysis of theso defined subsequent problems.

Technically, if further use is to be made of the Kalman filter in assessing the Columbia river problem, it would be desirable to investigate the validity of assuming Gaussian distributions for the measurement errors and other sources of uncertainty. Significantly skewed distributions could compromise interpretation of the robustness of the predicted policy outcomes. Likewise, if (deterministic) parametric uncertainties are to be "unpacked" from the single aggregate (of the matrix Σ_Q), and a filtering framework remains the preferred computational setting, this could be achieved through the relatively minor extension of the extended Kalman filter (as in Beck and Halfon (1991)).

ALTERNATIVES

The obvious alternative to using the Kalman filter on a problem of this nature is Monte Carlo simulation, or some variation on that theme. Had this alternative been adopted, uncertainty would almost certainly have been accounted for in a different manner. In particular, as with virtually all Monte Carlo studies, the uncertainty attaching to the (deterministic) parameters of the model would have been the sole source of uncertainty accounted for. The question for calibration would then have been that of using the past observed temperatures in order to constrain, in some way, the choice of candidate parameterisations to be used for predicting the outcomes of the policy alternatives. Normally, one encounters Monte Carlo simulation in the context of forecasting (not model calibration). This requires specification of the statistical distributions to be used for the model's parameters, treated as random variables. In the absence of past observations, ranges of parameter values drawn from the literature are used to define these distributions. It is unusual to find studies using the set of past observations to generate "posterior" distributions of the parameters, for the purpose of forecasting, with the calibration process started with the "prior", literature-derived distributions.

In short, we derive models from uncertain theories reconciled with uncertain observations; we make predictions that are uncertain using models whose uncertainties will reflect all the successes and failures of calibration; and we must make decisions that are robust in the face of the resulting uncertain predictions, i.e., we must determine whether we would opt for the same course of action, all the uncertainties notwithstanding. Conceptually, the Kalman filter fits well with this view. If the alternative of Monte Carlo simulation were to be considered, it would probably find appropriate implementation through the procedure of Generalised Likelihood Uncertainty Estimation (GLUE) of Beven and Binley (1992).

POSSIBLE ISSUES FOR FURTHER STUDY

In the light of what has just been stated, regarding the account taken of uncertainty, from model development, through calibration and forecasting, into decision-making, the following could be of some significance. If one accepts the suggestion that the critical decision will turn on the reliability of the forecasts of maximum temperatures, then the manner in which the model is calibrated -- as the instrument of making this particular prediction -- should be geared to this goal. In practical terms, this implies that the covariance-matching technique employed for choosing Σ_Q should seek the best possible match over the periods of the summer maxima (as opposed to other seasons of the year, or over the year in some average manner). Figures 22 through 29 of the Report do not fully illuminate whether such a strategy has been pursued. We may probably conclude it has not.

Two criteria are used separately to rank the three policy alternatives, the number of days during the year when the temperature standard is exceeded and the magnitude of the excess temperature. It may be more meaningful to discriminate on the basis of a composite criterion, designed to capture the sense that the joint action of duration and magnitude of the excess is vital for the well-being of the endangered fish.

The option of removing the dams, in spite of the express consideration of uncertainty, still promises to bring about a significant change in the status quo. This is apparent from Figure 34 (when compared Figure 33) and, marginally more so, from the comparison of Figures 39 and 40. Making decisions under uncertainty -- as opposed to the determinism prevailing in its absence -- introduces greater subtlety (and complexity) into the debate. For example, in another context (Klepper *et al.*, 1991) the consequence of an action was forecast to have the effect of increasing the mean value of a commercial mussel culture, but also of introducing (relative to the status quo) a non-negligible risk of population collapse. While it is apparent that the present Report could have sustained such a more elaborate discussion, none is provided.

CONCLUSIONS AND RECOMMENDATIONS

This EPA Report, in my opinion, should contribute beneficially to the debate surrounding the survival of endangered species of fish in the Columbia River, precisely because of the way in which it casts its analysis in the setting of uncertainty and risk.

Although an unusual method to use, the Kalman filter has been implemented in a technically sound manner. On the whole the approximations and assumptions made in this implementation are consistent with the style of the investigation, this being that of a screening analysis. By implication, therefore, further study is likely to be needed before decisions on managing the thermal regime of the Columbia and Snake Rivers can be made.

Clarification should be sought on the following points: (i) the precise nature of the "simulated" values plotted in Figures 6 through 13; (ii) the possible impact on the predicted results of the policy alternatives of the likely higher uncertainties attaching to the model's hydraulic parameters in the event of removing the dams; (iii) the possible significance of covariance (as opposed to variance) among the sources of uncertainty accounted for in Σ_Q; and (iv) the consistency of interpretation of the results shown in Figures 14 through 21 relative to Figures 6 through 13.

If further study is to be undertaken by the EPA, one should seek to have the following issues addressed (among others raised in this review):

(i) a sensitivity analysis of the influence on prediction uncertainty of (a) an enlarged system noise variance-covariance matrix (Σ_Q), as a consequence of removing the dams, and (b) an altered set of values for the elements of this matrix as a result of gearing its calibration to the goal of matching covariances for the summer temperature maxima;

(ii) an assessment of prediction uncertainty when the specific sources of uncertainty are separated out from the aggregated form of Σ_Q, with a view to ranking the relative importance of these different sources;

(iii) an assessment of the normality of the distributions of various quantities manipulated through the filtering algorithms;

(iv) a more elaborate treatment of the implications of these, and any similar, subsequent, results for the debate surrounding decision-making under uncertainty.

REFERENCES

Beck, M B, and Halfon, E (1991), "Uncertainty, Identifiability and the Propagation of Prediction Errors: A Case Study of Lake Ontario", *J Forecasting*, **10**(1&2), pp 135-161.

Beven, K J, and Binley, A M (1992), "The Future of Distributed Models: Model Calibration and Predictive Uncertainty", *Hydrological Processes*, **6**, pp 279-298.

Klepper, O, Scholten, H, and van der Kamer, J P G (1991), "Prediction Uncertainty in an Ecological Model of the Oosterschelde Estuary", *J Forecasting*, **10**(1&2), pp 191-209.

Mehra, R K (1972), "Approaches to Adaptive Filtering", *IEEE Transactions on Automatic Control*, **10**, pp 693-698.

Van Geer, F C, Te Stroet, C B M, and Zhou, Y (1991), "Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations. 1. The Role of System Noise and its Calibration", *Water Resources Research*, **27**(8), pp 1987-1994.

M B Beck

(b) (6)
Athens
Georgia 30606-5734